

(12) UK Patent Application (19) GB (11) 2 368 167 (13) A

(43) Date of A Publication 24.04.2002

(21) Application No 0025445.8

(22) Date of Filing 17.10.2000

(71) Applicant(s)
Radka Radana Dvorak
36 Epsom Road, GUILDFORD, Surrey, GU1 3LE,
United Kingdom

(72) Inventor(s)
Radka Radana Dvorak

(74) Agent and/or Address for Service
Reddie & Grose
16 Theobalds Road, LONDON, WC1X 8PL,
United Kingdom

(51) INT CL⁷
G06F 17/24

(52) UK CL (Edition T)
G4H HTAT H1A H13D H14A

(56) Documents Cited
EP 0280866 A2 EP 0250677 A1

(58) Field of Search
UK CL (Edition S) G4H HTAT
INT CL⁷ G06F
ONLINE:WPI,EPODOC,JAPIO

BEST AVAILABLE COPY

(54) Abstract Title
Knowledge management software system

(57) A document authoring tool uses reader centred linguistic techniques for producing new versions of electronic documents. The basic text and a topic list is imported and a rule set applied to analyse text to identify and define communicative functions. Links relating to a topic are identified and each topic assigned a communicative function to inform readers what the topic is about in relation to the text. The new document created displays active links as well as the communicative functions relating to the topics to which the links relate, allowing the users to navigate a document with knowledge about the purpose of the topic within the document.

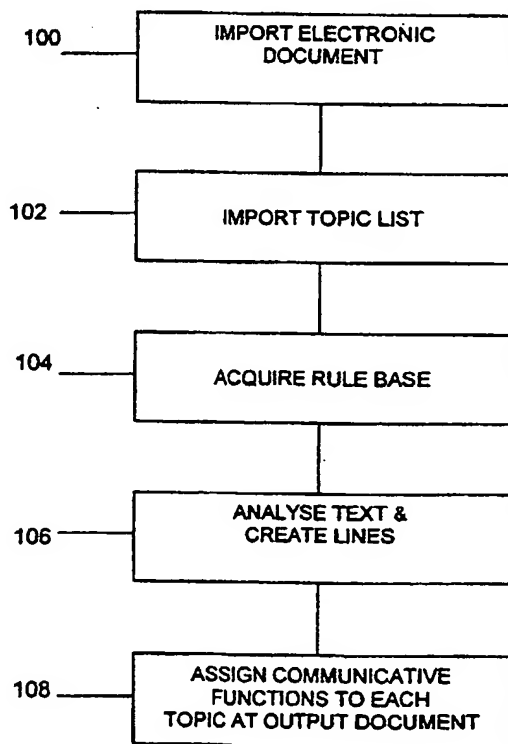


FIGURE 1

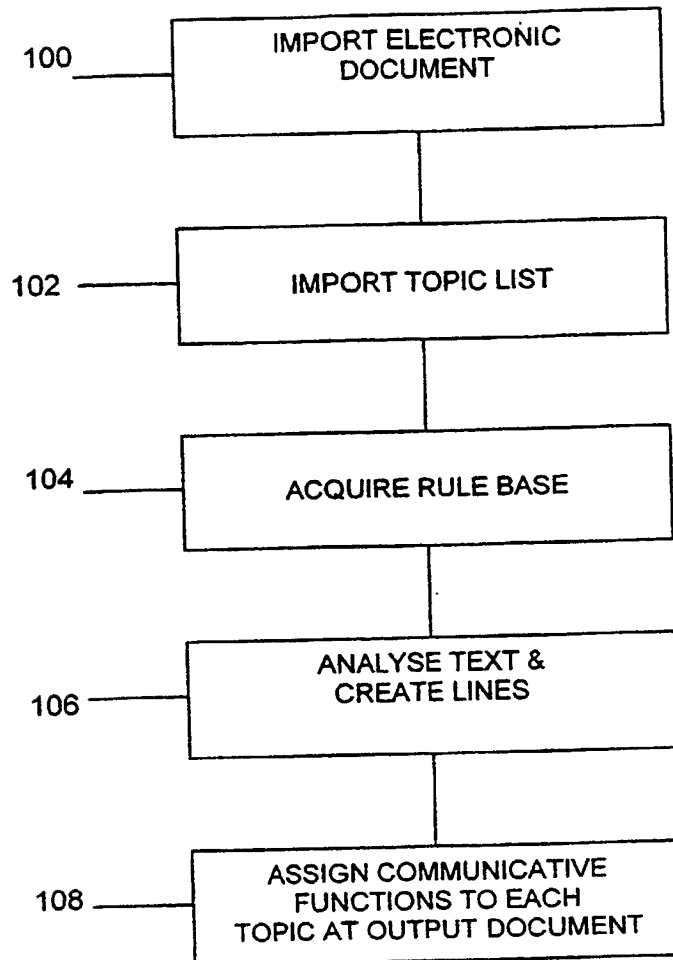


FIGURE 1

217

144 146 148 150 152

142

138

136

134

122

124

126

128

130

120

140

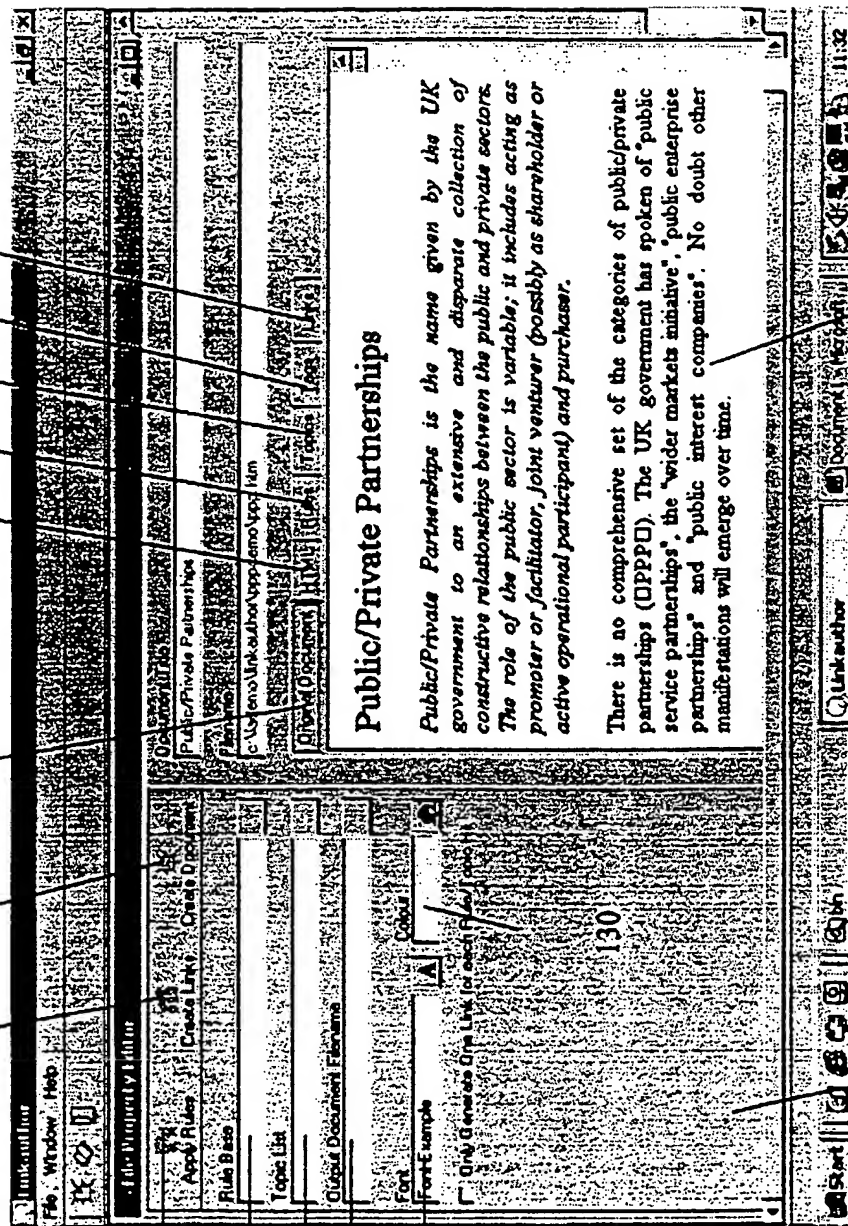
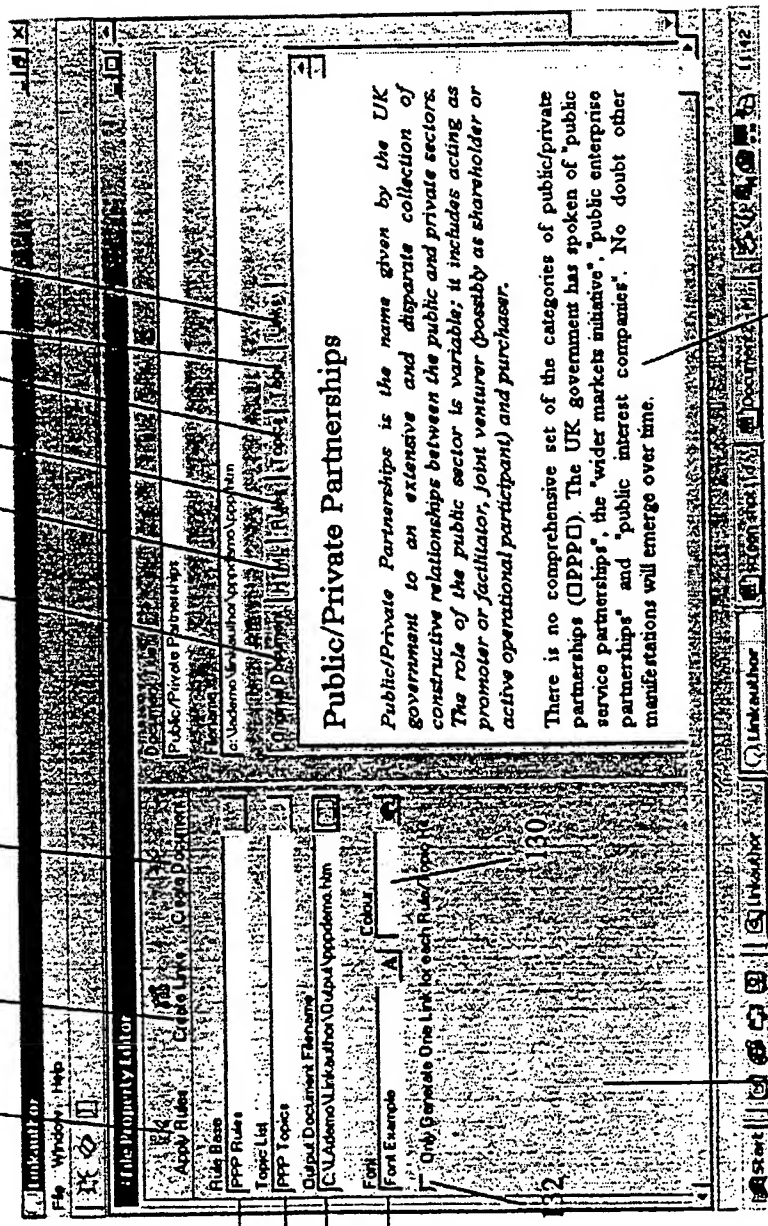


FIGURE 2

317
142 144 146 148 150 152

134 136 138



122
124
126
128

FIGURE 3

120 140

202403

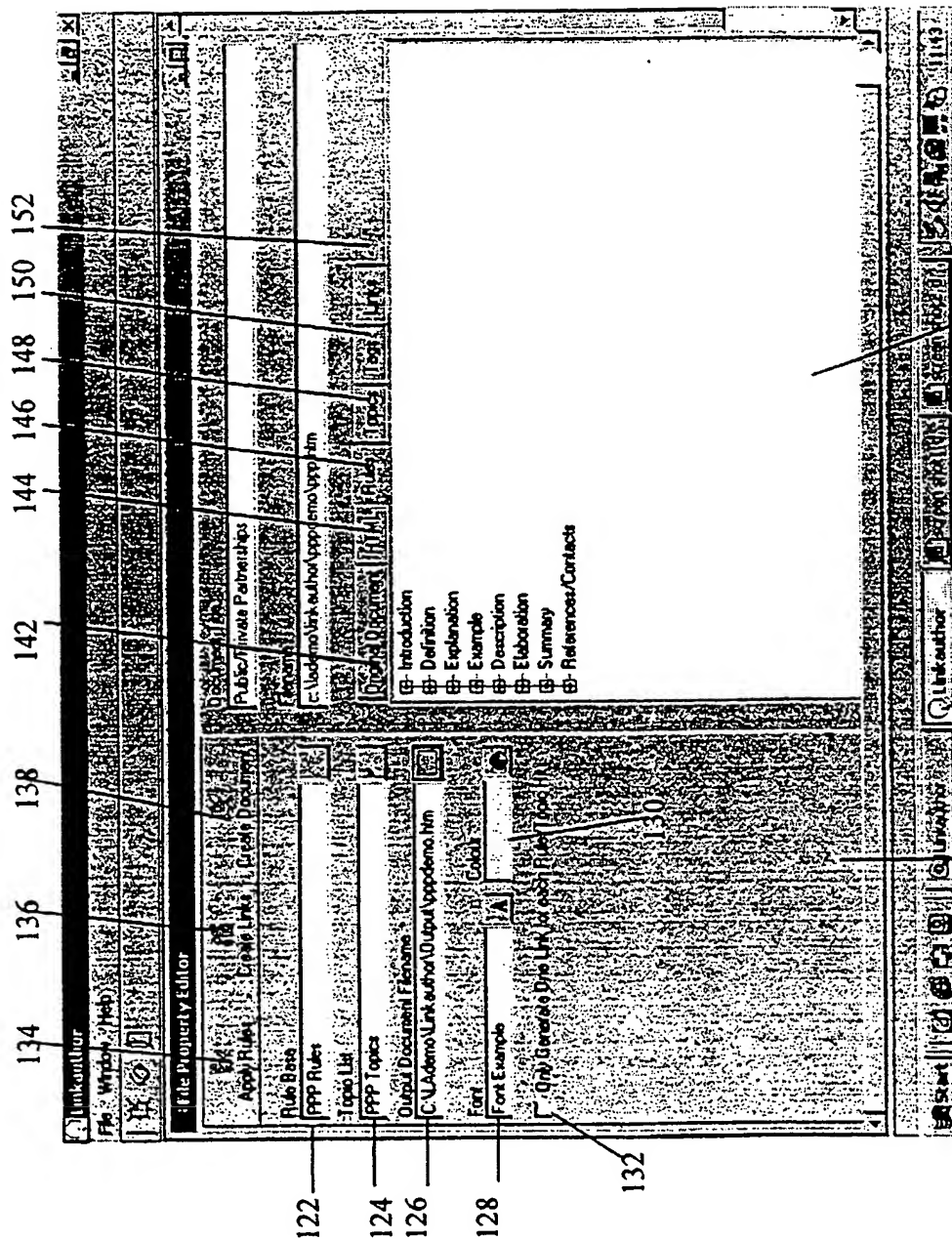


FIGURE 4

Linkauthor Generated Page - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Public Private Partnerships

PPP, Definition

Public Private Partnerships or PPP is the name given by the UK government to an extensive and disparate collection of constructive relationships between the public and private sectors. The role of the public sector is variable; it includes acting as promoter or facilitator, joint venturer (possibly as shareholder or active operational participant) and purchaser.

There is no comprehensive set of the categories of public/private partnerships. The UK government has spoken of "public services partnerships", the "wider markets initiative", "public enterprise partnerships" and "public interest companies". No doubt other manifestations will emerge over time.

The factors which have driven government to promote PPP include:

- state withdrawn from commercial activities, either for political or economic reasons
- a desire to improve the quality of public services
- a desire to obtain better value for money (which includes managing the public sector spend over the long term).

PPP, Explanation

PPP is seen as an altogether different concept from outright privatisation. It provides a structure which exploits some of the lessons learned from privatisation, which are said to include:

- privatisation was seen as too blunt and abrupt a technique; the slower, balanced process of PPP is felt to enhance government value and provide a long term continuing participation of government in functions of public interest

156

158

Linkauthor

- Generated Links
- Project Co
- Example
- Definition
- PFI
- Bankability
- Contact
- PPP
- Summary
- Introduction
- Explanation
- Definition

FIGURE 5



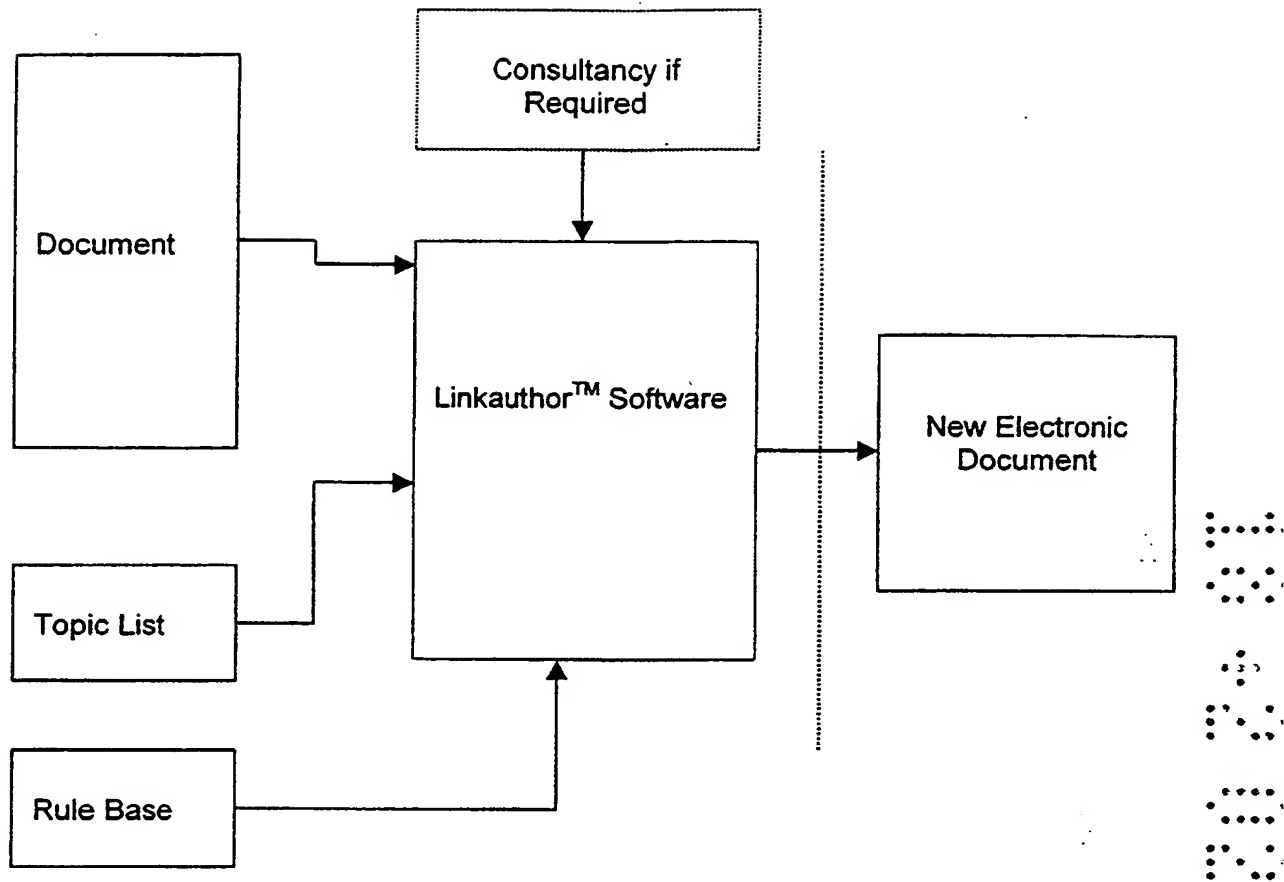
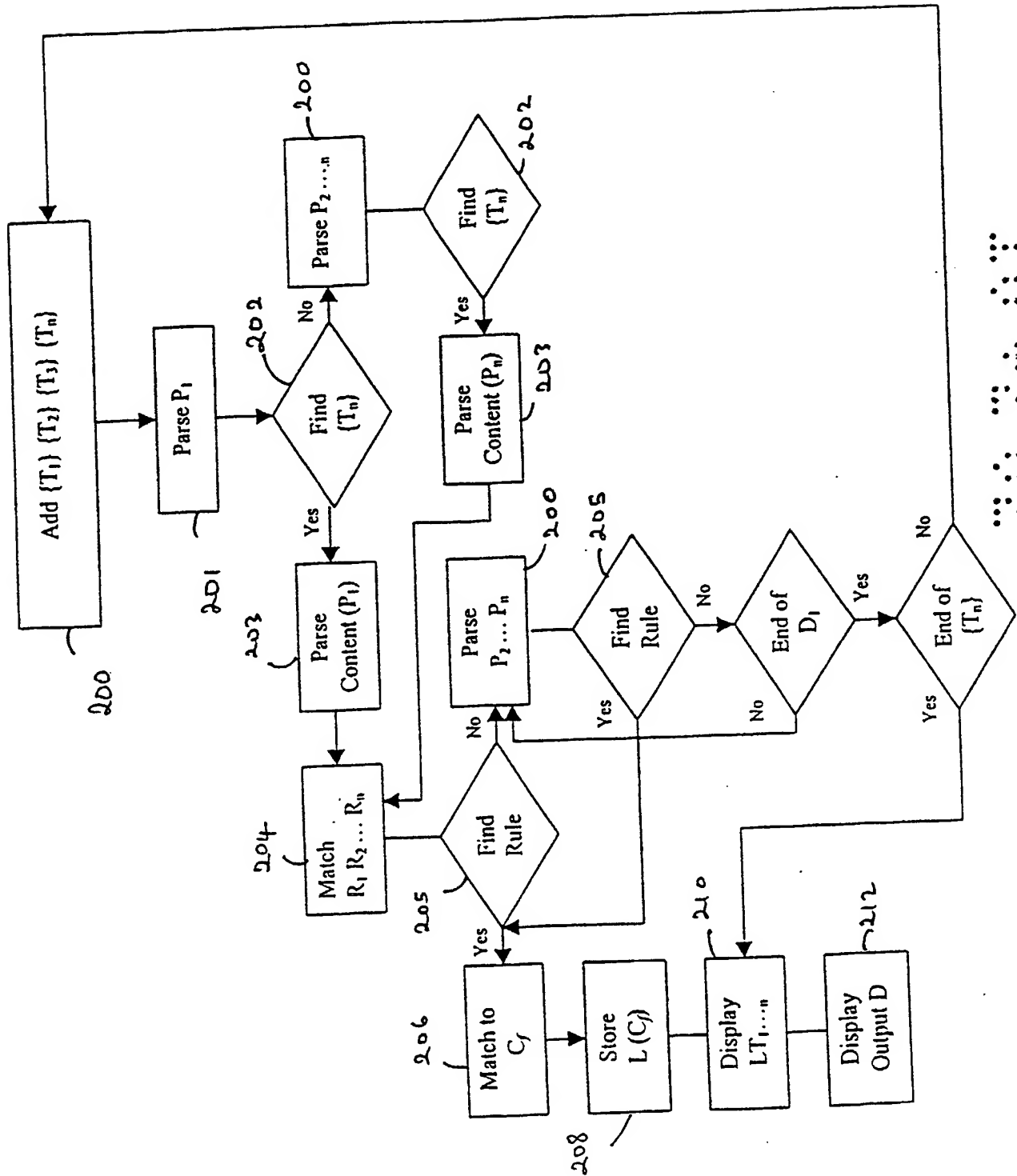


FIGURE 6



Key

(T) = Topic text
 P = Paragraph
 R = Rule
 Cf = Communication function
 LT = Link Topic
 D = Document

FIGURE 7

- 1 -

KNOWLEDGE MANAGEMENT SOFTWARE SYSTEM

This invention relates to knowledge management systems and in particular to authoring tools that electronically link information within and between documents.

5 It has been stated that one of the challenges facing information designers (hypertext/hypermedia) is to understand the user's specific task and to support links that follow from some model of the user's need for information in some particular context.

10 The advent of the world wide web and other computer networks has led to an explosion in the number of documents available in electronic form. Many of these documents, whether they are viewed in Netscape or Internet Explorer browser based web sites, Intranets, knowledge management documents,
15 electronic documents in CD ROM format or in some other form do not fulfill their true potential. Readers have to spend too much time looking for information and often get lost within the document structure. When hypertext links appear users are often compelled to follow them in case they are
20 led to relevant information. However, users often find themselves getting further and further away from their intended reading goals. Users often find these systems harder to use than traditional paper based equivalents.

There are known in the art, two main approaches for
25 assisting users to assimilate electronic documents. The first is author (text) centred in which the electronic document mimics the structure and organisation of the original document. This approach attempts to give the user or reader the same facilities as were available with the
30 paper based equivalent. The text is not restructured and a

key word search engine is added. This approach aims to make the paper based document more usable by delivering it in the electronic medium. Developers who have adopted this approach argue that changes should not be made to the original document as users have well developed models of the old medium. Examples of this approach can be found in products sold by Verity Inc. and Excalibur Technologies Inc.. The Verity product searches intelligently for information contained within documents held on companies internal networks. Successful searches generate a list of documents, ranked according to estimated relevance, for the use to select and read. The searches are based on concept based rules which allow the user to ask questions in familiar terms while increasing the relevancy of search results. The Excalibur product user Semantic Network and APRP (Adaptive Pattern Recognition Processing) to assist in finding relevant information in both corporate intranets and the World Wide Web. As in the Verity product, a list of documents matching user defined search criteria is created for the reader to select and read. When a document is selected, the text matching the search criteria is highlighted. The software selects one pieces of information as the most probable candidate for display and allows the user to navigate directly to it.

The second approach taken to document creation is system centred, usually influenced by a software tool, and using various metaphors or mathematically based formulae to organise and deliver the electronic text. Developers who create documents using these tools tend to create their own link structures using probability theorems and other mathematical based methodologies. An example of such an approach is seen in products provided by Autonomy, Inc. A

number of its products look for links within and between documents with a view to increasing the usability of information. However the theory underlying these products is based largely on mathematical theories such as Autonomy's
5 Bayesian mathematical probability theory rather than natural language theories. The products suffer from the disadvantage that they must be installed on all users' servers or PCs which increases costs and installation overheads.

10 Other known knowledge management systems are provided by Documentum Inc. and Multicosm Inc. Documentum provides a web based content management system which assists developers in delivering documents over the Internet and Intranets. The product is extensively used within the pharmaceutical
15 industry. The system only creates hypertext links through a tagging facility that is prescribed by site administrators. For example, customers, document types, and business roles can be used to create the linking. These links are not assigned degrees of importance through either
20 mathematical probability or natural language processing rules. The system is server based and suffers from the disadvantage of requiring a high level of training in order to set up and manage the content management engine and the web content management module. Extensive consultancy is
25 offered to system users which again generates additional cost.

The Multicosm product provides a knowledge management suite comprising a server, a databases and a number of agents. The system includes a dynamic link library organised by
30 themes. The links are retrieved by a reader as and when they are needed. The system suffers from a number of

disadvantages. Links have to be input by the developer and then checked by the creator of the document as each link is organised into a theme. The web master decides which themes should be kept, which should be deleted and which custom
5 links should be added. The final links are then stored in a linkbase. Control does not reside with the reader, but with the supply technical. Moreover, the system does not generate a read only document. It is used to apply links dynamically to all documents as they are served by the
10 browsers. As with many of the known systems, the product has to be installed on users' PCs or linked to a server before it can be used, creating a substantial cost and time overhead.

We have appreciated that none of the existing products are
15 wholly satisfactory and that there is a need for an improved knowledge management system which addresses the deficiencies in the existing systems.

A third approach to the creation of electronic documents has been proposed in 2 research papers: 'A methodology for user
20 centred link structures for textbook to hypertext conversion' published in Bruce D. Shrivvers (ed.) Proceedings of the 25th Hawaii International Conference on System Sciences. Pp 619-628. (1992) and *Using a hypertext
25 texbook: an evaluation of reader centered link structures'*, Proceedings of Computers and Writing '94 Conference 20-23 May, Columbia Missouri, USA. The approach is also described in a PhD Thesis of Radka R Dvorak, 1990-1994, awarded
December 1994 by Queen Mary and Westfield College, University of London, England and entitled 'Text to
30 Hypertext Books: an evaluative Investigation Into Reader Centred Structures For Hypertext'.

This third approach proposed a reader centred structure which focuses on the needs of the reader or user rather than the system itself or the structure defined by the author of the original document. Existing text based, author based
5 linking methodology cannot easily provide purposeful and meaningful links as the author or information engineer inevitably cannot predict the reader's prior knowledge of the topic, the reader's goals in manipulating electronic text nor their strategies for achieving these goals.

10 Reader-centred/user centred link structures are developed from linguistic theory which appeals to rhetorical predicated as a basic unit for the discourse strategies writers use to communicate information to their readers. Rhetorical predicates characterise the communicative
15 functionality in text-based discourse and form the basis for links between the text units. These are the link structures.

The invention in its broadest form provides an authoring tool which can analyse an electronic document according to
20 a set of rules which are reader-centred and identify communicative functions in the text. Links to an imported list of topics are determined and each topic assigned a communicative function. The result is a document that can be navigated by context.

25 More specifically there is provided a method of creating a fresh electronic document from an input electronic document, without altering the content of the original document, comprising the steps of: acquiring an input electronic document; acquiring a list of topics
30 relevant to the input electronic document; applying a set of

rules to analyse the input electronic document to identify communicative functions embodied in text comprising the input electronic document and create links in the input electronic document to the topics in the list of topics; assigning a communicative function to each of the identified topics; and generating the fresh electronic document comprising the input electronic document, the links and the communicative functions assigned to the topics to which the links relate, whereby the links are displayed with the communicative functions associated with the topics to which the links relate.

The invention also provides apparatus for creating a fresh electronic document from an input electronic document, without altering the content of the original document comprising: means for acquiring an input electronic document; means for acquiring a list of topics relevant to the input electronic document; a rule based engine for applying a set of rules to analyse the input electronic document to identify communicative functions embodied in document text comprising the input electronic document; means for creating links in the input electronic document to the topics in the list of topics; means for assigning a communicative function to each of the identified topics; and means for generating the fresh electronic document comprising the input electronic document, the links and the communicative functions assigned to the topics to which the links relate, whereby the links are displayed with the communicative functions associated with the topics to which the links relate.

Other preferred features of the invention are set out in the dependent claims to which reference should be made.

Embodiments of the invention will now be described, with reference to the accompanying drawings in which:

Figure 1 is a basic flow chart illustrating operating of an embodiment of the invention;

Figure 2 is a screen shot of a menu frame;

Figure 3 is a similar view to Figure 2 after the user
5 has selected topic lists and rule sets;

Figure 4 is a similar view to Figure 2 and 3 with the rule set shown in the viewing window;

Figure 5 is a screen shot of an output frame showing an output electronic document after operation of the invention;

10 Figure 6 is a schematic block diagram showing a system embodying the invention; and

Figure 7 is a flow chart illustrating the algorithm underlying the invention;

The information technology authoring tool to be described is
15 based on the reader centred approach to linking information in documents. It uses the rule based engine, the link structures, framework and methodology to confer documents from one platform to another. The reader/user centred approach is exploited by the authoring tool by applying two
20 link structures.

The first is called reader-centred (user-centred) primary structure. This segments text into units defined by discourse-theoretic notions of topicality and communicative function. The link structure is not made explicit to the
25 reader/user but is the engine of the tool that links topics within documents and between documents.

The second structure is called reader-centred (user-centred) secondary structure. This explicitly cues the reader/user to the communicative function in relation to a node (a
30 coherent chunk of text that can be defined by one or more paragraphs) topic that is served by the traversal of a link. It is this secondary structure that is made explicit to the reader/user. The communicative functions act as links,

which inform the users what the topic is about before they activate, that is traverse, a link.

5 The communicative functions have been identified by a number of linguists although the original credit goes back to Aristotle who first recorded the use of rhetorical predicates. He described examples and enthymemes as two of the means a speaker can use for a successful persuasive argument. Enthymemes are argument types using a form or reasoning where conclusions are deduced by having a common
10 proposition

The purpose of the authoring tool to be described is to assist individuals and companies in managing information by creating usable documents for various types of users within the target population for which they have been created, supporting readers differing information seeking strategies.
15 The tool has a novel high speed rule base engine embedded in it which is based on linguistic models. Whilst prior art exists which can assist users in finding topics and linking to the topic within large documents they find, none can help the reader locate the relevant topic they need to satisfy
20 their information goal. Indeed many existing tools have become so good at identifying relevant documents that they create an information overload for the reader, actually increasing the time required for analysis. Embodiments of the present invention address this problem by adding
25 semantic information to the key words in documents to inform the reader of what the topic is about before they have to start following links into the document.

Figure 1 shows a basic overview of the main steps performed
30 by the authoring tool. At 100, a basic electronic document is imported from an outside source. The document may be in any form, for example raw HTML data or in a word processed format. At step 102, a list of topics is imported. The topic list is a list of key words covered by the document.

The topics can be imported electronically, for example an index, table of contents, from an ASCII or Word document. Alternatively they may be inputted by the knowledge management administrator, the information designer or the tool user, or they may be generated electronically by other software as is known from the Verity Inc, system mentioned above. At step 104 a set of linguistic rules is imported. This is the rule base engine which defines how the tool processes the document.

At step 106, the tool automatically analyses the words and phrases that identify communicative functions and creates and stores links relating to a topic within the document. This will be described in greater detail.

At 108, the new electronic document is generated and displayed through an industry standard browser. A communication function is assigned to each topic (key word). These communicative functions are hyperlinks to specific sections of the document. Each topic is listed only once, though it may have any number of links. These links are able to support readers' information seeking goals because they are informing readers what function the topic is serving within certain sections of the document. Thus each topic serves some purpose and fulfils an information-seeking goal. a communicative function is assigned to each link. This informs the reader/user what the topic or key word is about in relation to the chunk of text. Thus each topic serves some purpose in completing an information seeking goal. Topics may serves the following purposes in a paragraph: define, explain, provide an example, illustrate, compare, introduce summarise, conclude, etc. This list is only exemplary. Thus navigation through the document is contextual based. It will be appreciated that by assigning communicative functions to each topic the reader can gain an

idea of the content before activating a link which avoids the pursuit of fruitless links.

5 The tool generates output files in a format which can be viewed in any conventional browser such as Netscape Navigator v.4.0 and above or Microsoft Internet Explorer v 4.0 and above. This has the advantage that the tool does not have to be loaded onto individual PCS which greatly reduces cost and installation time. Output documents may be fully customised in terms of layout, graphics or colours.
10 Other than local disk storage limitations there is no limit to the size of a document that can be processed.

The tool generates a default two frame set comprising a menu frame, examples of which are shown in figures 2 to 4, and a main viewing frame, and example of which is shown in figure
15 5. It will be appreciated that other frame set formats are possible and that the output files can be customised by knowledge management professional and information engineers. When a hypertext link is activated, the paragraph where the topic appears in relation to the communicative function is
20 highlighted in the main viewing area. End-users, therefore, are given the relevant information about the link before they have to click on the link and move to a particular section of the text within the document.

Referring now to figure 2 there is shown a screen shot of
25 the menu frame 120. On the left hand side are a series of windows which allow the user to select the rule base 122, the topic list 124, the file name and the destination of the output document 126, the font 128 and the colour 130. The user also has the option at 132 of only generating a single
30 link for each rule/topic hit. At the top left hand of the screen are three icons 134, 136, 138. The first of these is clicked by the user to apply the selected rulebase. The second 136 is clicked by the user to create links and the

third, 138 is clicked by the user to create the final authored document.

At the top right hand side of the screen are displayed the document title and the filename. Below the filename is a
5 window 140 in which is displayed the document to be authored. In the example shown this is entitled "Public Private Partnerships". The user can select from a menu above the window what is displayed. In this example the tab 142 entitled "original document" has been selected. Other
10 selectable tabs are HTML 144 which display the HTML code for the document, Rules tab 146 which displays the set of rules selected, Topics tab 148 which displays the topic list for the document, tags tab 150 which displays a list of tags for the document, and link tab 152 which displays the links,
15 once they have been generated.

Figure 3 shows an example of the same screen in which the Rulebase, Topic list and Output file have been selected. Figure 4 is similar to Figure 3 except that the create rules icon 136 has been clicked and so the list of rules is
20 displayed.

Figure 5 shows the final authored document. It will be seen that the textual content of the document is the same. However, under the heading Definition 154, which is highlighted as a link, the tool has selected the portion of
25 text which it has identified from its textual analysis as being a definition of the subject matter, Public Private Partnerships. The link however includes the topic so that the reader can tell to what the link relates. A second link 156 with the associated topic entitled Explanation is also
30 highlighted and leads the user to a body of text that explains what is a Public Private Partnership.

At the left hand side of the screen is displayed a list 158 of the links that have been generated. These are

hierarchical as can be seen from the link entitled PPP. There are four additional links, summary, introduction, explanation and definition. The use can see what each of these links will take him or her to before activating them.

5 The manner in which the tool functions will now be described in more detail with reference to figures 6 to 8. Figure 6 is a schematic block diagram which illustrates how the initial electronic document 160, the topic list, 162 and the rule base 164 are imported into the authoring tool 166, and
10 a new electronic document 170 is output.

Figure 7 is a flow chart showing how the process operates. A document has n paragraphs of text. At step 200, the topic text is added. The first paragraph is then parsed at step 201. At 202, the process asks whether any topic text has
15 been found. If none is found, the process return to step 200 and parses the next paragraph. If topic text is found, at step 203, the process parses the content of the paragraph and at 204, the process tries to match with each of the rules in the rule set. If no match is possible the process
20 again returns to step 200 and parses the next paragraph. At step 205 the process asks whether a matching rule has been found. If it the process proceeds to step 206. At this step a match is made to the communicative function. At step
25 208, the Link topics are stored. These are the highlighted links described with reference to figure 5 earlier. When the tool is processing the document, it doesn't export the links and topics to the browser (final document) as it is locating the links. It finds the topic, finds the communicative function, adds a number to the html, and
30 stores this. When the whole document is processed, then all the topics and links are exported to the left frame, see figure 6 and the document to the right frame. At step 210 the link topics are displayed and at step 212 the output document is displayed.

It will be appreciated that once the process has gone as far as it can for a paragraph, which may end up with link topics being created, the process will return to the beginning and operate on the next paragraph.

5 The algorithm operated by the tool is as follows:

1. Rule Parsing

Syntax: tx where t = text x = subscript

 { }y where { } = token y = subscript

10 1.1 General rule format with two tokens: {T}, {.}

1.2 Parser will parse rules of this form:

to {.}0t1 {.}1.....tn {.}n tn+1 {T}n+1 tn+2{.}n+2.....tm {.}m te

NOTE: One space between all text and tokens, and no spaces before or after text and tokens.

15 2. Parser

2.1. Replace {T} with current topic text

2.2 Start from beginning of paragraph

2.3 Search for to

 If found, start from location of t o

20 Search for ti

Search for tn
If found, start from location of tn

2.4 Concatenate tn+1, Topic Text, tn+2 and search

5 If found, start from location of tn+1, Topic, tn+2
Search for tn+3

10 Search for tm
If found, start from location of tm

Search for te
If found, RULE APPLIES

3. Browser & engine

15	Rule Base	Engine Topics	Browser Topics	Labels
	R1	T1	LT1	+LT1
	R2	T2	LT2	+LT1
	R3	T3	LT3	+LT1
	R4	T4		+Describing
		T5		-Elaborating
		T6		T1
		*****		+LTm

		*****n		

Engine Topics: String to search for in HTML document

Browser Topics: Which topics are displayed in Linkset frame
on browser

5 4. Assumptions

4.1 Every rule has one and only one Cf

4.2 All topics can have one or more Cfs

4.3 In general rule format, any tx or { }y can be null.

10 4.4 Each Cf one or more rules. Thus, the Cf is assigned to
the topic once a rule is identified by the parser.

4.5 Divide document into paragraphs, and parse per
paragraph.

5. Link types and Rules defined and described

15 Communicative functions are used explicitly at the browser
level. They are assigned by the rules.

Cf1=R1, R2, R3, R4

Cf2=R5, R6, R7, R8

Cf3=R9, R10, R11, R12

Cfn=Rn, Rn+1, Rn+2, Rn+y

20 R1=[(fx), T], R2=[(fx), T] R3=[(fx), T], R4=[(fx), T], R5=[(fx), T], Rn=[(fx), T].

f(x) = proposition

f=VP x=NP

VP= verb phrases

NP= noun phrases

Cf= identified communicative functions, = (rhetorical predicates).

5 **For example in label above:**

 'Elaborate' rule applied to T1, T2, T3 and T1, T2, T3 are grouped under LT3. So, the labels tab shows exactly what will be displayed in the left frame for the end users. The frame is interactive, it is a drop-down menu, generated using JavaScript. This functionality is not language specific.

 It will be seen from the foregoing that the tool acts as an intelligent and highly visual search engine which accurately identifies words and phrases in large volumes of documents. It analyses their usage and relevance to the end user's reading need and the message the author was trying to convey, and extracts topics with representative concepts (communicative function) that accurately identify the content of each document. Identification of topics and linking topics, analysing the use to which they are put in terms of reading needs and explicitly extracting these concepts are the most important functional aspects which enable a text management tool to produce a usable electronic document.

25 Embodiments of the invention described have many advantages. Firstly, the resulting new electronic documents that are created are more effectively navigable resulting in more efficient and productive use of user's time. Second, use of the tool removes the need for laborious and expensive detailed mark-up and editing of text. The tool is able to find the key words and generate the links without altering the original structure or chopping up the text to deliver

only pieces of information to the end-users. It delivers the whole document as it was created. The tool can also generate documents in industry standard format without the need to install additional software at user's PCs. This
5 results in a further considerable cost saving. Partly because user organisations do not have to support the tool over a potentially wide user base and partly because documents generated by the tool can be published as web pages accessible via the Internet.

10 The invention is not limited to the specific examples described. Many variations are possible and will occur to those skilled in the art. For example, the rule based engine can be used in a variety of types of application, whether or not they use a standard browser. The principles
15 of textual analysis are the same and can be used in various media or embedded in larger software packages.

The invention is limited only by the scope of the attached claims.

CLAIMS

1. A method of creating a fresh electronic document from an input electronic document, without altering the content of the original document, comprising the steps of:

5 acquiring an input electronic document;
 acquiring a list of topics relevant to the input electronic document;

 applying a set of rules to analyse the input electronic document to identify communicative functions embodied in text comprising the input electronic document and to create
10 links in the input electronic document and to create links in the input electronic document to the topics in the list of topics;

 assigning a communicative function to each of the identified topics; and

15 generating the fresh electronic document comprising the input electronic document, the links and the communicative functions assigned to the topics to which the links relate, whereby the links are displayed with the communicative functions associated with the topics to which the links
20 relate.

2. A method according to claim 1, wherein the links and associated communicative functions are displayed and can be highlighted with respect to the text of the input electronic document.

25 3. A method according to claim 1 or 2, wherein the rule base applies a set of reader-centred rules based on a linguistic theory.

4. A method according to claim 1, 2 or 3, wherein the step of applying the set of rules comprises the steps of:

30 parsing a paragraph of the input electronic document to identify topic text;

matching identified topic text with one of a plurality of rules; and

matching the topic text and rule with a communicative function.

5

5. A method according to any of claims 1 to 4, further comprising displaying the links and assigned communicative functions as a list separate from the fresh electronic document.

10

6. A method according to any of claims 1 to 5, comprising highlighting a section of text in which a topic appears when a link is activated which includes a communicative function assigned to that topic.

15

7. A method according to any preceding claim, wherein the fresh electronic document is viewable with an Internet browser or other software producing output files.

8. A method according to claim 7, wherein the fresh electronic document is displayed in a viewing frame.

20

9. A computer program which, when run on a computer, causes the computer to perform the method of any of claims 1 to 8.

25

10. A computer readable storage media for recorder thereon a computer program which, when run on a computer, causes the computer to perform the method of any of claims 1 to 8.

30

11. Apparatus for creating a fresh electronic document from an input electronic document, comprising:
means for acquiring an input electronic document;
means for acquiring a list of topics relevant to the
input electronic document;

a rule based engine for applying a set of rules to analyse the input electronic document to identify communicative functions embodied in document text comprising the input electronic document;

5 means for creating links in the input electronic document to the topics in the list of topics;

means for assigning a communicative function to each of the identified topics; and

10 means for generating the fresh electronic document comprising the input electronic document, the links and the communicative functions assigned to the topics to which the links relate, whereby the links are displayed with the communicative functions associated with the topics to which the links relate.

15 12. Apparatus according to claim 11, wherein the links and associated communicative functions are displayed and can be highlighted with respect to the text of the input electronic document.

20 13. Apparatus according to claim 11 or 12, wherein the rule base engine applies a set of reader-centred rules based on a linguistic theory.

14. Apparatus according to claim 11, 12 or 13, wherein the rule based engine rules comprises:

25 means for parsing a paragraph of the input electronic document to identify topic text;

means for matching identified topic text with one of a plurality of rules; and

means for matching the topic text and rule with a communicative function.

30

15. Apparatus according to any of claims 11 to 14, further comprising means for displaying the links and assigned communicative functions as a list separate from the fresh electronic document.

16. Apparatus of any of claims 11 to 15, comprising means for highlighting a section of text in which a topic appears when a link is activated which includes a communicative function assigned to that topic.

5 17. Apparatus according to any preceding claim, wherein the fresh electronic document is viewable with an Internet browser or other software producing output files.

18. Apparatus according to claim 17, wherein the fresh electronic document is displayed in a viewing frame.

10 19. A method of creating a fresh electronic document from an input electronic document, substantially as herein described with reference to the accompanying drawings.

20. Apparatus for creating a fresh electronic document from an input electronic document, substantially as herein
15 described with reference to the accompanying drawings.



Application No: GB 0025445.8
Claims searched: 1-20

Examiner: Mike Davis
Date of search: 8 January 2001

Patents Act 1977 Search Report under Section 17

Databases searched:

UK Patent Office collections, including GB, EP, WO & US patent specifications, in:

UK Cl (Ed.S): G4H (HTAT)

Int Cl (Ed.7): G06F

Other: Online: WPI, EPODOC, JAPIO

Documents considered to be relevant:

Category	Identity of document and relevant passage	Relevant to claims
A	EP 0280866 A2 (IBM)	-
A	EP 0250677 A1 (TOSHIBA)	-

X	Document indicating lack of novelty or inventive step	A	Document indicating technological background and/or state of the art.
Y	Document indicating lack of inventive step if combined with one or more other documents of same category.	P	Document published on or after the declared priority date but before the filing date of this invention.
&	Member of the same patent family	E	Patent document published on or after, but with priority date earlier than, the filing date of this application.

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ **BLACK BORDERS**
- ☐ **IMAGE CUT OFF AT TOP, BOTTOM OR SIDES**
- ☐ **FADED TEXT OR DRAWING**
- ☐ **BLURRED OR ILLEGIBLE TEXT OR DRAWING**
- ☐ **SKEWED/SLANTED IMAGES**
- ☐ **COLOR OR BLACK AND WHITE PHOTOGRAPHS**
- ☐ **GRAY SCALE DOCUMENTS**
- ☐ **LINES OR MARKS ON ORIGINAL DOCUMENT**
- ☐ **REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY**
- ☐ **OTHER:** _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

THIS PAGE BLANK (USPTO)